FATTY ACID BIOSYNTHETIC PATHWAY

FIG.1

Gene Sequence of *sdd17*, an Omega-3 Fatty Acid Desaturase Gene
from *Saprolegnia diclina* (ATCC 56851)

```
   1  ATGACTGAGG ATAAGACGAA GGTCGAGTTC CCGACGCTCA CGGAGCTCAA
  51  GCACTCGATC CCGAACGCGT GCTTTGAGTC GAACCTCGGC CTCTCGCTCT
 101  ACTACACGGC CCGCGCGATC TTCAACGCGT CGGCCTCGGC GGCGCTGCTC
 151  TACGCGGCGC GCTCGACGCC GTTCATTGCC GATAACGTTC TGCTCCACGC
 201  GCTCGTTTGC GCCACCTACA TCTACGTGCA GGGCGTCATC TTCTGGGGCT
 251  TCTTCACGGT CGGCCACGAC TGCGGCCACT CGGCCTTCTC GCGCTACCAC
 301  AGCGTCAACT TTATCATCGG CTGCATCATG CACTCTGCGA TTTTGACGCC
 351  GTTCGAGAGC TGGCGCGTGA CGCACCGCCA CCACCACAAG AACACGGGCA
 401  ACATTGATAA GGACGAGATC TTTTACCCGC ACCGGTCGGT CAAGGACCTC
 451  CAGGACGTGC GCCAATGGGT CTACACGCTC GGCGGTGCGT GGTTTGTCTA
 501  CTTGAAGGTC GGGTATGCCC CGCGCACGAT GAGCCACTTT GACCCGTGGG
 551  ACCCGCTCCT CCTTCGCCGC GCGTCGGCCG TCATCGTGTC GCTCGGCGTC
 601  TGGGCCGCCT TCTTCGCCGC GTACGCGTAC CTCACATACT CGCTCGGCTT
 651  TGCCGTCATG GGCCTCTACT ACTATGCGCC GCTCTTTGTC TTTGCTTCGT
 701  TCCTCGTCAT TACGACCTTC TTGCACCACA ACGACGAAGC GACGCCGTGG
 751  TACGGCGACT CGGAGTGGAC GTACGTCAAG GGCAACCTCT CGAGCGTCGA
 801  CCGCTCGTAC GGCGCGTTCG TGGACAACCT GAGCCACCAC ATTGGCACGC
 851  ACCAGGTCCA CCACTTGTTC CCGATCATTC CGCACTACAA GCTCAACGAA
 901  GCCACCAAGC ACTTTGCGGC CGCGTACCCG CACCTCGTGC GCAGGAACGA
 951  CGAGCCCATC ATCACGGCCT TCTTCAAGAC CGCGCACCTC TTTGTCAACT
1001  ACGGCGCTGT GCCCGAGACG GCGCAGATCT TCACGCTCAA AGAGTCGGCC
1051  GCGGCCGCCA AGGCCAAGTC GGACTAA
```

FIG.2

Amino Acid Sequence of an Omega-3 Fatty Acid Desaturase (SDD17) from
Saprolegnia diclina (ATCC 56851)

```
  1  MTEDKTKVEF PTLTELKHSI PNACFESNLG LSLYYTARAI FNASASAALL

 51  YAARSTPFIA DNVLLHALVC ATYIYVQGVI FWGFFTVGHD CGHSAFSRYH

101  SVNFIIGCIM HSAILTPFES WRVTHRHHHK NTGNIDKDEI FYPHRSVKDL

151  QDVRQWVYTL GGAWFVYLKV GYAPRTMSHF DPWDPLLLRR ASAVIVSLGV

201  WAAFFAAYAY LTYSLGFAVM GLYYYAPLFV FASFLVITTF LHHNDEATPW

251  YGDSEWTYVK GNLSSVDRSY GAFVDNLSHH IGTHQVHHLF PIIPHYKLNE

301  ATKHFAAAYP HLVRRNDEPI ITAFFKTAHL FVNYGAVPET AQIFTLKESA

351  AAAKAKSD*
```

FIG.3

Comparative analysis of *S. diclina* Delta 17-desaturase (SDD17.pep)
& *Synechocystis* sp. Delta 15-desaturase (SYCDESB)

Frame: 2 initn: 733 init1: 305 opt: 689 Z-score: 996.8 expect(): 1.5e-47
  40.9% identity in 269 aa overlap
(76-336:204-471)

```
                50        60        70        80        90       100
SDD17.pep    SAALLYAARSTPFIADNVLLHALVCATYIYVQGVIFWGFFTVGHDCGHSAFSRYHSVNFI
                         :||::||::|:||||||||::||: :::|
SYCDESB      YFFLDVGLIAGFYALAAYLDSWFFYPIFWLIQGTLFWSLFVVGHDCGHGSFSKSKTLNNW
                530       560       590       620       650       680


                110       120       130       140       150       160
SDD17.pep    IGCIMHSAILTPFESWRVTHRHHHKNTGNIDKDEIFYPHRSVKDLQ----DVRQWVYTLG
             || : |: ||:|:::||::|| || |||||| || :||    |  |    :     |·
SYCDESB      IGHLSHTPILVPYHGWRISHRTHHANTGNIDTDESWYPVSEQKYNQMAWYEKLLRFYLPL
                710       740       770       800       830       860


                170       180       190       200 .     210       220
SDD17.pep    GAWFVYLKVGYAPRTMSHFDPWDPLLL-RRASAVIVSLGVWAAFFAAYAYLTYSLGFAVM
             |: :||     |   ||| | :||:   : :||::|  : |||  :  ::||:::|:   :
SYCDESB      IAYPIYLFRRSPNRQGSHFMPGSPLFRPGEKAAVLTSTFALAAFVGFLGFLTWQFGWLFL
                890       920       950       980      1010      1040


                230       240       250       260       270       280
SDD17.pep    GLYYYAPLFVFASFLVITTFLHHNDEATPWYGDSEWTYVKGNLSSVDRSYGAFVDNLSHH
             :| || :||: :|   ::|||||::: ||| ::| ::||  ||::||:||  |:: : |
SYCDESB      LKFYVAPYLVFVVWLDLVTFLHHTEDNIPWYRGDDWYFLKGALSTIDRDYG-FINPIHHD
               1070      1100      1130      1160      1190      1220


                290       300       310       320       330
SDD17.pep    IGTHQVHHLFPIIPHYKLNEATKHFAAAYPHLVRRNDEPIITAFFKT---AHLFVNYGAV
             |||| :||:|  :|||||  :||| :  : | :|||| ||||:   |:  :  | |:
SYCDESB      IGTHVAHHIFSNMPHYKLRRATEAIKPILGEYYRYSDEPIWQAFFKSYWACHFVPNQGSG
               1250      1280      1310      1340      1370      1400


             340       350
SDD17.pep    PETAQIFTLKESAAAAKAKSD

SYCDESB      VYYQSPSNGGYQKKPXLILIESNQHREGRQYXMVLLPSDRLMRSMEEVKQSHSKRSALNQ
               1430      1460      1490      1520      1550      1580
```

FIG.4

Comparative analysis of *S. diclina* Delta 17-desaturase (SDD17.pep)
C. elegans Delta 17-desaturase (CELEFAT)

Frame: 1 initn: 490 init1: 222 opt: 502 Z-score: 724.0 expect(): 2.3e-32
 31.6% identity in 310 aa overlap
 (2-303:49-347)

```
                                      10        20        30
                             MTEDKTKVEFPTLTELKHSIPNACFESNLGL
SDD17.pep                    |  :: ::::||:  ::::||  ||| :|
CELFAT       VTGGDVLVDARASLEEKEAPRDVNANTKQATTEEPRIQLPTVDAFRRAIPAHCFERDLVK
                 80        110       140       170       200       230


                40        50        60        70        80        90
SDD17.pep    SLYYTARAIFNASASAALLYAARSTPFIADNVLLHALVCATYIYVQGVIFWGFFTVGHDC
             |:||   |||      :  | : |:  ||  ::  ||: :::|:|||||
CELFAT       SIRYLVQDF------AALTILYFALPAFEYFGLFGYLVWNIFM---GVFGFALFVVGHDC
                 260       290       320       350       380


                100       110       120       130       140
SDD17.pep    GHSAFSRYHSVNFIIGCIMHSAILTPFESWRVTHRHHHKNTGNIDKD--EIFYPHRSVKD
             |::||   ::| :|||  | |:::|:  |: :|:  || |::||||  :::   :: :
CELFAT       LHGSFSDNQNLNDFIGHIAFSPLFSPYFPWQKSHKLHHAFTNHIDKDHGHVWIQDKDWEA
                 410       440       470       500       530       560


                150       160       170       180       190       200
SDD17.pep      LQDVRQWV----YTLGGAWF-VYLKVGYAPRTMSHFDPWDPLLLRRASAVIVSLGVWAAF
             : : ::|      ::    || || |:    |||  |::  |::| :: |   ::
CELFAT       MPSWKRWFNPIPFSGWLKWFPVYTLFGFC--DGSHFWPYSSLFVRNSDRVQCVISGICCC
                 590       620       650       680       710       740


                210       220       230       240       250       260
SDD17.pep    FAAYAYLTYSLGFAVMGLYYYAPLFVFASFLVITTFLHHNDEATPWYGDSEWTYVKGNLS
             ||  || ||: :::   ||::||  |: :||:|:|:|  |:::   | :||::|:|: :
CELFAT       VCAYIALTIAGSYSNWFWYYWVPLSFFGLMLVIVTYLQHVDDVAEVYEADEWSFVRGQTQ
                 770       800       830       860       890       920


                270       280       290       300       310       320
SDD17.pep    SVDRSYGAFVDNLSHHI-GTHQVHHLFPIIPHYKLNEATKHFAAAYPHLVRRNDEPIITA
             ::|| ||  :|:  ||| ||| |  :||:| ||||:| |||: |||
CELFAT       TIDRYYGLGLDTTMHHITDGHVAHHFFNKIPHYHLIEATEGVKKVLEPLSDTQYGYKSQV
                 950       980       1010      1040      1070      1100


                330       340       350
SDD17.pep    FFKTAHLFVNYGAVPETAQIFTLKESAAAAKAKSD
CELFAT       NYDFFARFLWFNYKLDYLVHKTAGIMQFRTTLEEKAKAKAXKNIPCRSRVQQQLLRFHRFC
                 1130      1160      1190      1220      1250      1280
```

FIG.5

Gene Sequence of *sdd12*, a Delta 12-Desaturase Gene
from *Saprolegnia diclina* (ATCC 56851)

```
   1  ATGTGCAAAG GTCAAGCTCC TTCCAAGGCC GACGTGTTCC ACGCTGCGGG
  51  GTACCGCCCG GTCGCCGGCA CGCCCGAGCC GCTGCCGCTG GAGCCCCCGA
 101  CGATCACGCT CAAGGACCTG CGCGCGGCGA TCCCGGCCCA CTGCTTTGAG
 151  CGCAGCGCTG CCACTAGCTT TTACCATTTG GCCAAGAACC TTGCGATCTG
 201  CGCCGGCGTG TTCGCCGTTG GCCTCAAGCT CGCGGCTGCC GACTTGCCGC
 251  TCGCGGCCAA GCTGGTCGCG TGGCCCATCT ACTGGTTCGT CCAGGGCACG
 301  TACTTTACGG GCATCTGGGT CATTGCGCAC GAATGCGGCC ACCAGGCGTT
 351  CTCGGCGTCC GAGATCCTCA ACGACACGGT CGGTATCATT CTTCACTCGC
 401  TCCTCTTTGT GCCGTACCAC AGCTGGAAGA TCACGCACCG CCGCCACCAC
 451  TCCAACACGG GCAGCTGCGA GAACGACGAG GTGTTTACGC CGACGCCGCG
 501  GTCCGTCGTC GAGGCCAAGC ACGACCACTC GCTCCTCGAA GAGAGCCCGC
 551  TCTACAACCT GTACGGCATC GTCATGATGC TTCTCGTGGG CTGGATGCCG
 601  GGCTACCTCT TCTTCAACGC GACCGGCCCG ACCAAGTACG CTGGCCTCGC
 651  CAAGTCGCAC TTCAACCCGT ACGCAGCCTT TTTCCTCCCA AAGGAGCGCC
 701  TCAGCATCTG GTGGAGCGAC CTCTGCTTCC TCGCGGCCTT GTACGGCTTT
 751  GGCTACGGCG TCTCGGTCTT CGGCCTCCTC GATGTCGCCC GCCACTACAT
 801  CGTGCCGTAC CTCATTTGCA ACGCGTACCT CGTGCTCATC ACGTACCTCC
 851  AGCACACGGA TACGTACGTG CCCCACTTCC GCGGCGACGA GTGGAACTGG
 901  CTGCGCGGCG CGCTCTGCAC CGTCGACCGC TCGTTCGGCG CGTGGATCGA
 951  CAGCGCGATC CACCACATTG CCGACACGCA CGTGACGCAC CACATTTTCT
1001  CCAAGACGCC CTTCTACCAC GCGATCGAGG CGACCGACGC CATCACGCCC
1051  CTCCTCGGCA AGTACTACCT CATCGACCCG ACGCCGATCC CGCTGGCGCT
1101  CTGGCGCTCG TTCACGCACT GCAAGTACGT CGAGGACGAC GGCAACGTTG
1151  TGTTTTACAA GCGCAAGCTC GAGGAAAAGT AA
```

FIG.6

Amino Acid Sequence of a Delta 12-Desaturase (SDD12) from *Saprolegnia diclina* (ATCC 56851)

```
  1  MCKGQAPSKA DVFHAAGYRP VAGTPEPLPL EPPTITLKDL RAAIPAHCFE

 51  RSAATSFYHL AKNLAICAGV FAVGLKLAAA DLPLAAKLVA WPIYWFVQGT

101  YFTGIWVIAH ECGHQAFSAS EILNDTVGII LHSLLFVPYH SWKITHRRHH

151  SNTGSCENDE VFTPTPRSVV EAKHDHSLLE ESPLYNLYGI VMMLLVGWMP

201  GYLFFNATGP TKYAGLAKSH FNPYAAFFLP KERLSIWWSD LCFLAALYGF

251  GYGVSVFGLL DVARHYIVPY LICNAYLVLI TYLQHTDTYV PHFRGDEWNW

301  LRGALCTVDR SFGAWIDSAI HHIADTHVTH HIFSKTPFYH AIEATDAITP

351  LLGKYYLIDP TPIPLALWRS FTHCKYVEDD GNVVFYKRKL EEK*
```

FIG.7

Comparative analysis of *S. diclina* Delta 12-desaturase (SDD12.pep)
& *G. hirsutum* Delta 12-desaturase (GHO6DES)

Frame: 3 initn: 992 init1: 413 opt: 1086 Z-score: 1531.8 expect(): 2.3e-77
45.6% identity in 379 aa overlap
(9-380:14-384)

```
                  10        20              30        40
SDD12.pep     MCKGQAPSKADVFHAAGYRPVAGTPEP------LPLEPPTITLKDLRAAIPAHC
                  ||: : |:| | :| |           :|:| | :|| ::: ||| ||
GHO6DESAT     LRVSSTWRXTAFFKASKMGAGGRMPIDGIKEENRGSVNRVPIEKPPFTLGQIKQAIPPHC
                  10        40        70        100       130       160

              50        60        70        80        90        100
SDD12.pep     FERSAATSFYHLAKNLAICAGVFAVGLKLAAADLPLAAKLVAWPIYWFVQGTYFTGIWVI
              |:||   || :::::| :  : :: :     ||   : :|||:|| :||  :||:|||
GHO6DESAT     FRRSLLRSFSYVVHDLCLASFFYYIATSYFHF-LPQPFSYIAWPVYWVLQGCILTGVWVI
              190       220       250       280       310       340

              110       120       130       140       150       160
SDD12.pep     AHECGHQAFSASEILNDTVGIILHSLLFVPYHSWKITHRRHHSNTGSCENDEVFTPTPRS
              ||| ||:|| : ::||||:|||| |:||| |||| :||||||||||| | ||||:| |:|
GHO6DESAT     AHEWGHHAFRDYQWVDDTVGLILHSALLVPYFSWKISHRRHHSNTGSMERDEVFVPKPKS
              370       400       430       460       490       520

              170       180       190       200       210       220
SDD12.pep     VVEAKHDHSLLEESPLYNLYGIVMMLLVGWMPGYLFFNATGPTKYAGLAKSHFNPYAAFF
              :       :     ::|  :  ::|: |:|| | | || ||::| :|   ||:||| ::
GHO6DESAT     KLSC---FAKYLNNPPGRVLSLVVTLTLGW-PMYLAFNVSG--RYYDRLASHYNPYGPIY
                    550       580       610       640       670

              230       240       250       260       270       280
SDD12.pep     LPKERLSIWWSDLCFLAALYGFGYGVSVFGLLDVARHYIVPYLICNAYLVLITYLQHTDT
              :|||::: || ::|::|  :   ::: ||  |::| ||:|||||||||||||||||||
GHO6DESAT     SDRERLQVYISDTGIFAVIYVLYKIAATKGLAWLLCTYGVPLLIVNAFLVLITYLQHTHS
              700       730       760       790       820       850

              290       300       310       320       330       340
SDD12.pep     YVPHFRGDEWNWLRGALCTVDRSFGAWIDSAIHHIADTHVTHHIFSKTPFYHAIEATDAI
              :||:  ::||:|||||||: ||:|| | :::::|:|:||||:||:|| | |||:||| ||
GHO6DESAT     ALPHYDSSEWDWLRGALSTMDRDFGV-LNKVFHNITDTHVAHHLFSTMPHYHAMEATKAI
              880       910       940       970       1000      1030

              350       360       370       380       390
SDD12.pep     TPLLGKYYLIDPTPIPLALWRSFTHCKYVEDDGNVVFYKRKLEEK
              | :||||| :| ||| |:|| :| |||| ||||
GHO6DESAT     KPILGKYYPFDGTPIYKAMWREAKECLYVEPDVGGGGGGGSKGVFWYRNKFXRPTNCLIAG
              1060      1090      1120      1150      1180      1210
```

FIG.8

Sequence ID:

Sequence ID 1
5'-ATC CGC GCC GCC ATC CCC AAG CAC TGC TGG GTC AAG-3'

Sequence ID 2
5'- GCC CTC TTC GTC CTC GGC CAY GAC TGC GGC CAY GGC TCG TTC TCG-3'

Sequence ID 3
5'-GAG RTG GTA RTG GGG GAT CTG GGG GAA GAR RTG RTG GRY GAC RTG-3'

Sequence ID 4
5'-CCC TAC CAY GGC TGG CGC ATC TCG CAY CGC ACC CAY CAY CAG AAC-3'

Sequence ID 5
5'-GTT CTG RTG RTG GGT CCG RTG CGA GAT GCG CCA GCC RTG GTA GGG-3'

Sequence ID 6
5'- GGC TCG CAC TTC SAC CCC KAC TCG GAC CTC TTC GTC-3'

Sequence ID 7
5'- GAC GAA GAG GTC CGA GTM GGG GTW GAA GTG CGA GCC-3''

Sequence ID 8
5'- GCG CTG GAK GGT GGT GAG GCC GCC GCG GAW GSA CGA CCA-3'

Sequence ID 9
5'- CTG GGG GAA GAG RTG RTG GAT GAC RTG GGT GCC GAT GTC RTG RTG-3'

Sequence ID 10
5'- GGT GGC CTC GAY GAG RTG GTA RTG GGG GAT CTK GGG GAA GAR RTG-3'

Sequence ID 11
5'-GAG RTG GTA RTG GGG GAT CTG GGG GAA GAR RTG RTG GRY GAC RTG-3'

Sequence ID 12
5'-TAC GCG TAC CTC ACG TAC TCG CTC G-3'

Sequence ID 13
5'-TTC TTG CAC CAC AAC GAC GAA GCG ACG-3'

Sequence ID 14
5'-GGA GTG GAC GTA CGT CAA GGG CAA C-3'

Sequence ID 15
5'-TCA AGG GCA ACC TCT CGA GCG TCG AC-3'

Sequence ID 16
5'-CCC AGT CAC GAC GTT GTA AAA CGA CGG CCA G-3'

FIG.9A

Sequence ID 17
5'- AGC GGA TAA CAA TTT CAC ACA GGA AAC AGC -3'

Sequence ID 18
5'-GGT AAA AGA TCT CGT CCT TGT CGA TGT TGC-3'

Sequence ID 19
5'-GTC AAA GTG GCT CAT CGT GC-3'

Sequence ID 20
5'-CGA GCG AGT ACG TGA GGT ACG CGT AC-3'

Sequence ID 21
5'-TCA ACA GAA TTC ATG ACC GAG GAT AAG ACG AAG GTC GAG TTC CCG-3'

Sequence ID 22
5'-AAA AGA AAG CTT CGC TTC CTA GTC TTA GTC CGA CTT GGC CTT GGC-3'

Sequence ID 23
5'-TCA ACA AAG CTT ATG ACC GAG GAT AAG ACG AAG GTC GAG TTC CCG-3'

Sequence ID 24
5'-AAA AGA GAA TTC CGC TTC CTA GTC TTA GTC CGA CTT GGC CTT GGC-3'

Sequence ID 25

```
   1  ATGACTGAGG ATAAGACGAA GGTCGAGTTC CCGACGCTCA CGGAGCTCAA
  51  GCACTCGATC CCGAACGCGT GCTTTGAGTC GAACCTCGGC CTCTCGCTCT
 101  ACTACACGGC CCGCGCGATC TTCAACGCGT CGGCCTCGGC GGCGCTGCTC
 151  TACGCGGCGC GCTCGACGCC GTTCATTGCC GATAACGTTC TGCTCCACGC
 201  GCTCGTTTGC GCCACCTACA TCTACGTGCA GGGCGTCATC TTCTGGGGCT
 251  TCTTCACGGT CGGCCACGAC TGCGGCCACT CGGCCTTCTC GCGCTACCAC
 301  AGCGTCAACT TTATCATCGG CTGCATCATG CACTCTGCGA TTTTGACGCC
 351  GTTCGAGAGC TGGCGCGTGA CGCACCGCCA CCACCACAAG AACACGGGCA
 401  ACATTGATAA GGACGAGATC TTTTACCCGC ACCGGTCGGT CAAGGACCTC
 451  CAGGACGTGC GCCAATGGGT CTACACGCTC GGCGGTGCGT GGTTTGTCTA
 501  CTTGAAGGTC GGGTATGCCC CGCGCACGAT GAGCCACTTT GACCCGTGGG
 551  ACCCGCTCCT CCTTCGCCGC GCGTCGGCCG TCATCGTGTC GCTCGGCGTC
 601  TGGGCCGCCT TCTTCGCCGC GTACGCGTAC CTCACATACT CGCTCGGCTT
 651  TGCCGTCATG GGCCTCTACT ACTATGCGCC GCTCTTTGTC TTTGCTTCGT
 701  TCCTCGTCAT TACGACCTTC TTGCACCACA ACGACGAAGC GACGCCGTGG
 751  TACGGCGACT CGGAGTGGAC GTACGTCAAG GGCAACCTCT CGAGCGTCGA
 801  CCGCTCGTAC GGCGCGTTCG TGGACAACCT GAGCCACCAC ATTGGCACGC
 851  ACCAGGTCCA CCACTTGTTC CCGATCATTC CGCACTACAA GCTCAACGAA
 901  GCCACCAAGC ACTTTGCGGC CGCGTACCCG CACCTCGTGC GCAGGAACGA
 951  CGAGCCCATC ATCACGGCCT TCTTCAAGAC CGCGCACCTC TTTGTCAACT
1001  ACGGCGCTGT GCCCGAGACG GCGCAGATCT TCACGCTCAA AGAGTCGGCC
1051  GCGGCCGCCA AGGCCAAGTC GGACTAA
```

FIG.9B

Sequence ID 26

```
  1 MTEDKTKVEF PTLTELKHSI PNACFESNLG LSLYYTARAI FNASASAALL

 51 YAARSTPFIA DNVLLHALVC ATYIYVQGVI FWGFFTVGHD CGHSAFSRYH

101 SVNFIIGCIM HSAILTPFES WRVTHRHHHK NTGNIDKDEI FYPHRSVKDL

151 QDVRQWVYTL GGAWFVYLKV GYAPRTMSHF DPWDPLLLRR ASAVIVSLGV

201 WAAFFAAYAY LTYSLGFAVM GLYYYAPLFV FASFLVITTF LHHNDEATPW

251 YGDSEWTYVK GNLSSVDRSY GAFVDNLSHH IGTHQVHHLF PIIPHYKLNE

301 ATKHFAAAYP HLVRRNDEPI ITAFFKTAHL FVNYGAVPET AQIFTLKESA

351 AAAKAKSD*
```

Sequence ID 27

```
   1 ATGGCCCCGC AGACGGAGCT CCGCCAGCGC CACGCCGCCG TCGCCGAGAC
  51 GCCGGTGGCC GGCAAGAAGG CCTTTACATG GCAGGAGGTC GCGCAGCACA
 101 ACACGGCGGC CTCGGCCTGG ATCATTATCC GCGGCAAGGT CTACGACGTG
 151 ACCGAGTGGG CCAACAAGCA CCCCGGCGGC CGCGAGATGG TGCTGCTGCA
 201 CGCCGGTCGC GAGGCCACCG ACACGTTCGA CTCGTACCAC CCGTTCAGCG
 251 ACAAGGCCGA GTCGATCTTG AACAAGTATG AGATTGGCAC GTTCACGGGC
 301 CCGTCCGAGT TTCCGACCTT CAAGCCGGAC ACGGGCTTCT ACAAGGAGTG
 351 CCGCAAGCGC GTTGGCGAGT ACTTCAAGAA GAACAACCTC CATCCGCAGG
 401 ACGGCTTCCC GGGCCTCTGG CGCATGATGG TCGTGTTTGC GGTCGCCGGC
 451 CTCGCCTTGT ACGGCATGCA CTTTTCGACT ATCTTTGCGC TGCAGCTCGC
 501 GGCCGCGGCG CTCTTTGGCG TCTGCCAGGC GCTGCCGCTG CTCCACGTCA
 551 TGCACGACTC GTCGCACGCG TCGTACACCA ACATGCCGTT CTTCCATTAC
 601 GTCGTCGGCC GCTTTGCCAT GGACTGGTTT GCCGGCGGCT CGATGGTGTC
 651 ATGGCTCAAC CAGCACGTCG TGGGCCACCA CATCTACACG AACGTCGCGG
 701 GCTCGGACCC GGATCTTCCG GTCAACATGG ACGGCGACAT CCGCCGCATC
 751 GTGAACCGCC AGGTGTTCCA GCCCATGTAC GCATTCCAGC ACATCTACCT
 801 TCCGCCGCTC TATGGCGTGC TTGGCCTCAA GTTCCGCATC CAGGACTTCA
 851 CCGACACGTT CGGCTCGCAC ACGAACGGCC CGATCCGCGT CAACCCGCAC
 901 GCGCTCTCGA CGTGGATGGC CATGATCAGC TCCAAGTCGT CTGGGCCTT
 951 CTACCGCGTG TACCTTCCGC TTGCCGTGCT CCAGATGCCC ATCAAGACGT
1001 ACCTTGCGAT CTTCTTCCTC GCCGAGTTTG TCACGGGCTG GTACCTCGCG
1051 TTCAACTTCC AAGTAAGCCA TGTCTCGACC GAGTGCGGCT ACCCATGCGG
1101 CGACGAGGCC AAGATGGCGC TCCAGGACGA GTGGGCAGTC TCGCAGGTCA
1151 AGACGTCGGT CGACTACGCC ATGGCTCGT GGATGACGAC GTTCCTTGCC
1201 GGCGCGCTCA ACTACCAGGT CGTGCACCAC TTGTTCCCCA GCGTGTCGCA
1251 GTACCACTAC CCGGCGATCG CGCCCATCAT CGTCGACGTC TGCAAGGAGT
1301 ACAACATCAA GTACGCCATC TTGCCGGACT TTACGGCGGC GTTCGTTGCC
1351 CACTTGAAGC ACCTCCGCAA CATGGGCCAG CAGGGCATCG CCGCCACGAT
1401 CCACATGGGC TAA
```

FIG.9C

Sequence ID 28

```
  1 ATGGCAAACA GCAGCGTGTG GGATGATGTG GTGGGCCGCG TGGAGACCGG
 51 CGTGGACCAG TGGATGGATG GCGCCAAGCC GTACGCACTC ACCGATGGGC
101 TCCCGATGAT GGACGTGTCC ACCATGCTGG CATTCGAGGT GGGATACATG
151 GCCATGCTGC TCTTCGGCAT CCCGATCATG AAGCAGATGG AGAAGCCTTT
201 TGAGCTCAAG ACCATCAAGC TCTTGCACAA CTTGTTTCTC TTCGGACTTT
251 CCTTGTACAT GTGCGTGGAG ACCATCCGCC AGGCTATCCT CGGAGGCTAC
301 AAAGTGTTTG AAACGACAT GGAGAAGGGC AACGAGTCTC ATGCTCAGGG
351 CATGTCTCGC ATCGTGTACG TGTTCTGCGT GTCCAAGGCA TACGAGTTCT
401 TGGATACCGC CATCATGATC CTTTGCAAGA AGTTCAACCA GGTTTCCTTC
451 TTGCATGTGT ACCACCATGC CACCATTTTT GCCATCTGGT GGGCTATCGC
501 CAAGTACGCT CCAGGAGGTG ATGCGTACTT TTCAGTGATC CTCAACTCTT
551 TCGTGCACAC CGTCATGTAC GCATACTACT TCTTCTCCTC CCAAGGGTTC
601 GGGTTCGTGA AGCCAATCAA GCCGTACATC ACCACCCTTC AGATGACCCA
651 GTTCATGGCA ATGCTTGTGC AGTCCTTGTA CGACTACCTC TTCCCATGCG
701 ACTACCCACA GGCTCTTGTG CAGCTTCTTG GAGTGTACAT GATCACCTTG
751 CTTGCCCTCT TCGGCAACTT TTTTGTGCAG AGCTATCTTA AAAAGCCAAA
801 AAAGAGCAAG ACCAACTAA
```

Sequence ID 29

```
  1 MTVGFDETVT MDTVRNHNMP DDAWCAIHGT VYDITKFSKV HPGGDIIMLA

 51 AGKEATILFE TYHIKGVPDA VLRKYKVGKL PQGKKGETSH MPTGLDSASY

101 YSWDSEFYRV LRERVAKKLA EPGLMQRARM ELWAKAIFLL AGFWGSLYAM

151 CVLDPHGGAM VAAVTLGVFA AFVGTCIQHD GSHGAFSKSR FMNKAAGWTL

201 DMIGASAMTW EMQHVLGHHP YTNLIEMENG LAKVKGADVD PKKVDQESDP

251 DVFSTYPMLR LHPWHRQRFY HKFQHLYAPL IFGFMTINKV ISQDVGVVLR

301 KRLFQIDANC RYGSPWNVAR FWIMKLLTTL YMVALPMYMQ GPAQGLKLFF

351 MAHFTCGEVL ATMFIVNHII EGVSYASKDA VKGVMAPPRT VHGVTPMQVT

401 QKALSAAEST KSDADKTTMI PLNDWAAVQC QTSVNWAVGS WFWNHFSGGL

451 NHQIEHHCFP QNPHTVNVYI SGIVKETCEE YGVPYQAEIS LFSAYFKMLS

501 HLRTLGNEDL TAWST*
```

Sequence ID 30
5'-CCG SAG TTC ACS ATC AAG GAG ATC CGC GAS KSC ATC CCG GCC CAC TGC TTC -3'

Sequence ID 31
5'-GRS CTT CTT GAK GTG GWM SGT GGC CTC CTC GGC GTG GTA GWR CGG CAT-3'

Sequence ID 32
5'-CCS STC TAC TGG GCC TGC CAG GGT RTC GTC CTC ACS GGT GTC TGG-3'

FIG.9D

Sequence ID 33
5'-CCS STC TAC TGG ATC RYS CAG GGT RTC GTC KGY ACS GGT GTC TGG-3'

Sequence ID 34
5'-GGC GTG GTA GTG CGG CAT SMM CGA GAA GAR GTG GTG GGC GAC GTG-3'

Sequence ID 35
5'-CAC GTA CCT CCA GCA CAC GGA CAC CTA CG-3'

Sequence ID 36
5'- GAT CGA CAG CGC GAT CCA CCA CAT TGC-3'

Sequence ID 37
5'- CAA ATG GTA AAA GCT AGT GGC AGC GCT GC-3'

Sequence ID 38
5'-AGT ACG TGC CCT GGA CGA ACC AGT AGA TG-3'

Sequence ID 39
5'- TCA ACA GAA TTC ATG TGC AAA GGT CAA GCT CCT TCC AAG GCC GAC GTG -3'

Sequence ID 40
5'- AAA AGA AAG CTT TTA CTT TTC CTC GAG CTT GCG CTT GTA AAA CAC AAC-3'

Sequence ID 41
```
   1  ATGTGCAAAG GTCAAGCTCC TTCCAAGGCC GACGTGTTCC ACGCTGCGGG
  51  GTACCGCCCG GTCGCCGGCA CGCCCGAGCC GCTGCCGCTG GAGCCCCCGA
 101  CGATCACGCT CAAGGACCTG CGCGCGGCGA TCCCGGCCCA CTGCTTTGAG
 151  CGCAGCGCTG CCACTAGCTT TTACCATTTG GCCAAGAACC TTGCGATCTG
 201  CGCCGGCGTG TTCGCCGTTG GCCTCAAGCT CGCGGCTGCC GACTTGCCGC
 251  TCGCGGCCAA GCTGGTCGCG TGGCCCATCT ACTGGTTCGT CCAGGGCACG
 301  TACTTTACGG GCATCTGGGT CATTGCGCAC GAATGCGGCC ACCAGGCGTT
 351  CTCGGCGTCC GAGATCCTCA ACGACACGGT CGGTATCATT CTTCACTCGC
 401  TCCTCTTTGT GCCGTACCAC AGCTGGAAGA TCACGCACCG CCGCCACCAC
 451  TCCAACACGG GCAGCTGCGA GAACGACGAG GTGTTTACGC CGACGCCGCG
 501  GTCCGTCGTC GAGGCCAAGC ACGACCACTC GCTCCTCGAA GAGAGCCCGC
 551  TCTACAACCT GTACGGCATC GTCATGATGC TTCTCGTGGG CTGGATGCCG
 601  GGCTACCTCT TCTTCAACGC GACCGGCCCG ACCAAGTACG CTGGCCTCGC
 651  CAAGTCGCAC TTCAACCCGT ACGCAGCCTT TTTCCTCCCA AAGGAGCGCC
 701  TCAGCATCTG GTGGAGCGAC CTCTGCTTCC TCGCGGCCTT GTACGGCTTT
 751  GGCTACGGCG TCTCGGTCTT CGGCCTCCTC GATGTCGCCC GCCACTACAT
 801  CGTGCCGTAC CTCATTTGCA ACGCGTACCT CGTGCTCATC ACGTACCTCC
 851  AGCACACGGA TACGTACGTG CCCCACTTCC GCGGCGACGA GTGGAACTGG
 901  CTGCGCGGCG CGCTCTGCAC CGTCGACCGC TCGTTCGGCG CGTGGATCGA
 951  CAGCGCGATC CACCACATTG CCGACACGCA CGTGACGCAC CACATTTTCT
1001  CCAAGACGCC CTTCTACCAC GCGATCGAGG CGACCGACGC CATCACGCCC
1051  CTCCTCGGCA AGTACTACCT CATCGACCCG ACGCCGATCC CGCTGGCGCT
1101  CTGGCGCTCG TTCACGCACT GCAAGTACGT CGAGGACGAC GGCAACGTTG
1151  TGTTTTACAA GCGCAAGCTC GAGGAAAAGT AA
```

FIG.9E

Sequence ID 42
```
  1   MCKGQAPSKA  DVFHAAGYRP  VAGTPEPLPL  EPPTITLKDL  RAAIPAHCFE

 51   RSAATSFYHL  AKNLAICAGV  FAVGLKLAAA  DLPLAAKLVA  WPIYWFVQGT

101   YFTGIWVIAH  ECGHQAFSAS  EILNDTVGII  LHSLLFVPYH  SWKITHRRHH

151   SNTGSCENDE  VFTPTPRSVV  EAKHDHSLLE  ESPLYNLYGI  VMMLLVGWMP

201   GYLFFNATGP  TKYAGLAKSH  FNPYAAFFLP  KERLSIWWSD  LCFLAALYGF

251   GYGVSVFGLL  DVARHYIVPY  LICNAYLVLI  TYLQHTDTYV  PHFRGDEWNW

301   LRGALCTVDR  SFGAWIDSAI  HHIADTHVTH  HIFSKTPFYH  AIEATDAITP

351   LLGKYYLIDP  TPIPLALWRS  FTHCKYVEDD  GNVVFYKRKL  EEK*
```

FIG.9F